SARS-CoV-2 Codon Usage Bias Downregulates Host Expressed Genes With Similar Codon Usage

Andres Mariano Alonso [1,2] and Luis Diambra [2,3*]

[1] InTech, Universidad Nacional de San Martin, Chascomús, Argentina, [2] CONICET, Chascomús, Argentina, [3] CREG, Universidad Nacional de La Plata, La Plata, Argentina

Severe acute respiratory syndrome has spread quickly throughout the world and was declared a pandemic by the World Health Organization (WHO). The pathogenic agent is a new coronavirus (SARS-CoV-2) that infects pulmonary cells with great effectiveness. In this study we focus on the codon composition for the viral protein synthesis and its relationship with the protein synthesis of the host. Our analysis reveals that SARS-CoV-2 preferred codons have poor representation of G or C nucleotides in the third position, a characteristic which could result in an unbalance in the tRNAs pools of the infected cells with serious implications in host protein synthesis. By integrating this observation with proteomic data from infected cells, we observe a reduced translation rate of host proteins associated with highly expressed genes and that they share the codon usage bias of the virus. The functional analysis of these genes suggests that this mechanism of epistasis can contribute to understanding how this virus evades the immune response and the etiology of some deleterious collateral effect as a result of the viral replication. In this manner, our finding contributes to the understanding of the SARS-CoV-2 pathogeny and could be useful for the design of a vaccine based on the live attenuated strategy.

Keywords: SARS-CoV-2, codon usage bias, codon optimality, translational control, pathogeny, vaccine design

1. INTRODUCTION

The new SARS-CoV-2 coronavirus is the causative agent of the current pandemic of COVID-19. This highly pathogenic virus has quickly become the latest threat to modern life. From the end of 2019 and up to the redaction of this paper, this virus has infected over 17 million people, leading to mild symptoms of fever and to lung function reduction, severe acute respiratory syndrome (SARS), and even death. The current global death toll stands at 670,000; however, we are still far away from determining the final mortality figure. In the absence of vaccines or effective antiviral treatments against SARS-CoV-2, it is important to understand how this virus appropriates the host translation apparatus and subverts the immune defenses of infected cells. This can be the first step in the development of novel therapeutics.

As intracellular parasites, virus replication depends on the translational machinery of their cellular hosts to translate viral transcripts. Thus, virus replication requires ribosomes, tRNA, and translation factors from the cell host. On the other hand, codon usage bias is a feature of natural selection and affects the genomes of all domains of life. It is known that more frequently used codons are used for coding highly abundant proteins (Pan et al., 1998; Dana and Tuller, 2014; Quax et al., 2015; Diambra, 2017). Virus genomes also have preferences in the codon usage, but, in this

case, the bias is constrained by the host translational machinery (Shackelton et al., 2006). The effects of codon composition of a transcript on its translation have been reported in literature (Gingold and Pilpel, 2011; Plotkin and Kudla, 2011; Shah and Gilchrist, 2011; McCarthy et al., 2017) and is considered an important determinant of gene expression (Zhou et al., 2009; Tuller et al., 2010a,b). However, the codon usage of a gene can also affect the translation of other genes (Frumkin et al., 2018). In fact, the virus replication demands not only ribosomes but also a lot tRNA resources for the codons highest in demand (Chen et al., 2020). The consumption of specific tRNAs for the virus replication could thus be an alternative to controlling host protein synthesis machinery as well as generating deleterious collateral effects on the function of the host cells. Recent evidence supports the idea that high codon usage similarity between virus and host can lead to a deleterious effect on the host (Chen et al., 2020). Furthermore, Chen et al. (2020) have shown that codon composition of highly expressed viral genes regulates the tRNA availability, affecting the decoding time of codons in the infected cells. In this manner the viral infection can convert abundant codons into scarce codons, reshaping the human codon optimality pattern. In this sense, it has been proposed that dysregulation of the tRNA pool can lead to disease (Dhindsa et al., 2020). In fact, there is growing evidence supporting that synonymous variations in a coding sequence can affect its folding and/or resulting expression level (Tsai et al., 2008; Hunt et al., 2014; Buhr et al., 2016; Kirchner et al., 2017; von Herrmann et al., 2018; Dhindsa et al., 2020; Kim et al., 2020).

It is known that coronavirus genomes are poor in GC content ($\sim$ 40%) and that there exist a preferential use of A-ended or U-ended codons in these genomes (Gu et al., 2004; Kandeel et al., 2020). The last bias is usually characterized in the literature as the GC3 content. Indeed, the analysis performed by Gu et al. (2004) suggests that this compositional constraint is a major determinant to synonymous codon usage. Here, we study recent proteomic data from SARS-CoV-2 infected cells (Bojkova et al., 2020) from a novel point of view. By considering the codon usage of the virus ORFeome, we characterize a set of genes whose expression could be affected by the massive demand of the tRNA which implies the virus replication. We find that those host genes encoding proteins with similar codons to the virus ORFeome have a lower translational rate. Extrapolating this finding to highly expressed genes in lung, we find a small set of genes that can be downregulated. These genes are involved in translation, immune systems, and cell calcification, to name a few, and their roles in the SARS-CoV-2 cell pathogenesis should be the target of further studies.

2. MATERIALS AND METHODS

The coding sequences associated with the genome of SARS-CoV and SARS-CoV-2 were obtained from the NCBI (NC_004718.3 and NC_045512.2, respectively). Highly expressed genes in the lung and arterial tissues were retrieved from the GTEx portal (gtexportal.org). Proteomic and translatomic data from infected CACO-2 cells were retrieved from Supplementary

Material of Bojkova et al. (2020), which are publicly available (ProteomeXchange repository, ID = PXD017710). These data contain protein levels from 6,381 proteins and translational rate from 2,715 proteins, and the proteomic information is consequently not available for many proteins.
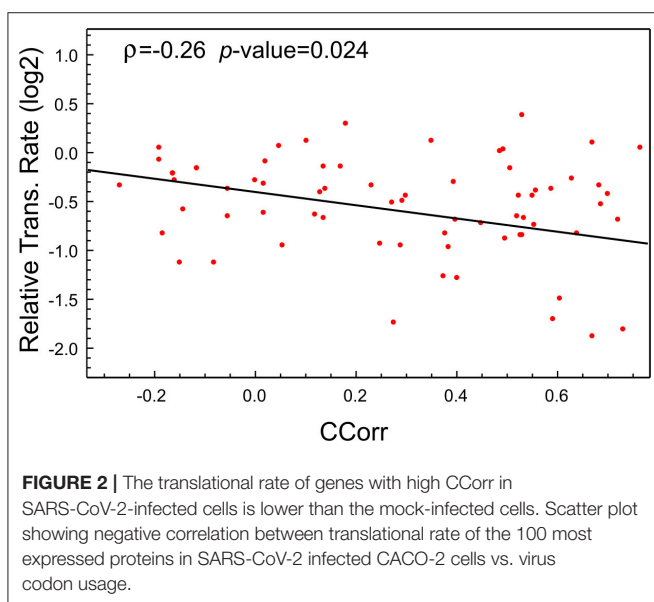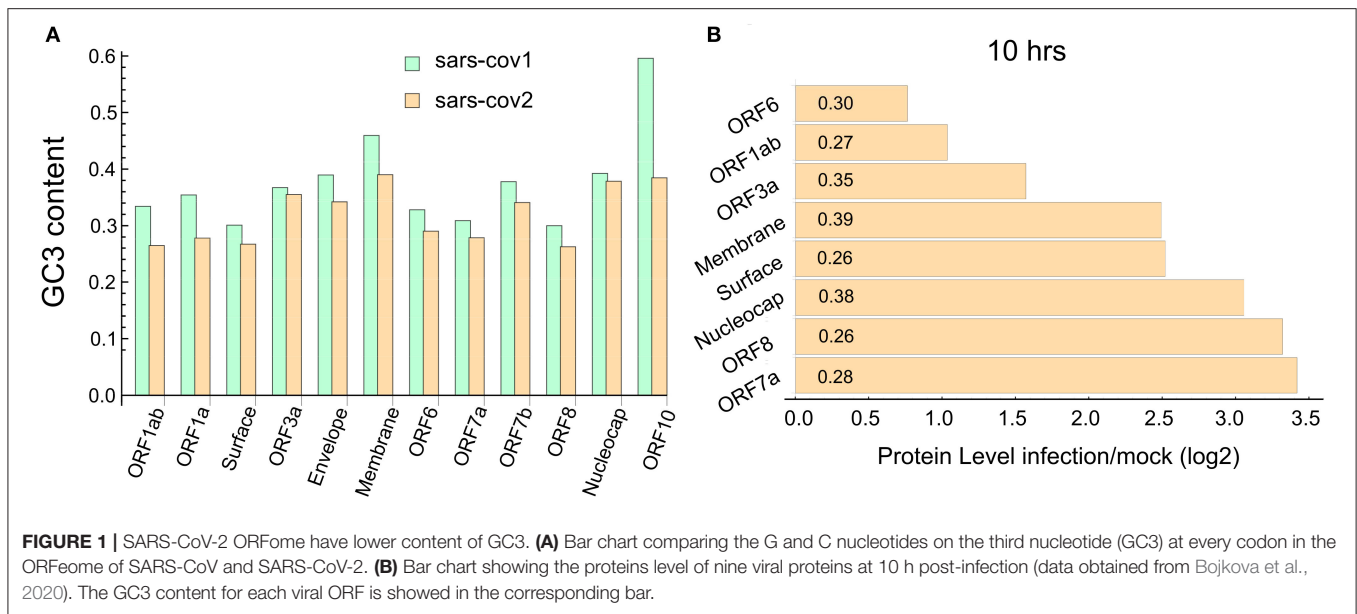
Given the coding sequence s, we compute the codon usage frequency as $f_s(c) = \frac{N_c}{L_s}$, where L_s is total number of codons in the sequence s, and N_c is the number of times that codon c is present in s. In a similar manner, we define codon usage frequency relative to the viral ORFome: $f_O(c) = \frac{N_c}{L_O}$, where L_O is the total number of codons in the ORFome, and N_c is the number of occurrences observed codon c in the ORFome. Thus, $f_s(c)$ and $f_O(c)$ are vector of 64 elements. To compute the CCorr of a given sequence s with viral ORFome, we consider the Pearson's correlation coefficient between the vector $f_s(c)$ and $f_O(c)$:

$$\text{CCorr} = \frac{\sum_c \left(f_s(c) - \overline{f_s}\right) \cdot \left(f_O(c) - \overline{f_O}\right)}{\sqrt{\sum_c \left(f_s(c) - \overline{f_s}\right)^2} \cdot \sqrt{\sum_c \left(f_O(c) - \overline{f_O}\right)^2}} \quad (1)$$

3. RESULTS

We study the codon composition of two complete coronavirus genomes by measuring the percentage of G and C nucleotides at the wobble position of the codons (GC3 content). **Figure 1A** depicts the GC3 content of each annotated coding sequence of the SARS-CoV and SARS-CoV-2. This analysis reveals that, in agreement with Gu et al. (2004) and Kandeel et al. (2020), codons preferentially used by these viruses are associated with a lower GC3 content than that expected from the random use of nucleotides. In particular, the ORFs corresponding to the proteins ORF1ab, ORF1a, surface, ORF6, ORF7a, and ORF8 exhibit a very low content of GC3. Specifically, the last two proteins are highly expressed at 10 h post-infection (PI), as can be seen in **Figure 1B**.

As was highlighted previously, this high expression of viral proteins could lead to an imbalance in the tRNA pool needed for the normal synthesis of the proteins of the host cell. Of course, as tRNA pools depend on the cell types, the postulated imbalance could be a tissue-dependent feature. To check this hypothesis, we make use of the recently available data about the proteome profile and translational rate in SARS-CoV-2-infected cells (Bojkova et al., 2020). In this study, a CACO-2 cell line was infected with SARS-CoV-2 and mocked, and the last one was used as the control. From this proteome, we selected the coding sequences corresponding to the 100 most abundant proteins in the mock infected cells at 10 h PI. We then computed the codon correlation (CCorr) between the codon usage of each one of these sequences and the codon usage of all coding sequences in SARS-CoV-2. **Figure 2** depicts a scatter plot of the translational rate of these genes from SARS-CoV-2-infected cells vs. the corresponding CCorr (red dots). The black line is the adjusted linear model, which shows a significant and negative correlation between the translation rate and the codon composition of each sequence. This means that the translational rate of coding

FIGURE 1 | SARS-CoV-2 ORFome have lower content of GC3. **(A)** Bar chart comparing the G and C nucleotides on the third nucleotide (GC3) at every codon in the ORFeome of SARS-CoV and SARS-CoV-2. **(B)** Bar chart showing the proteins level of nine viral proteins at 10 h post-infection (data obtained from Bojkova et al., 2020). The GC3 content for each viral ORF is showed in the corresponding bar.

FIGURE 2 | The translational rate of genes with high CCorr in SARS-CoV-2-infected cells is lower than the mock-infected cells. Scatter plot showing negative correlation between translational rate of the 100 most expressed proteins in SARS-CoV-2 infected CACO-2 cells vs. virus codon usage.
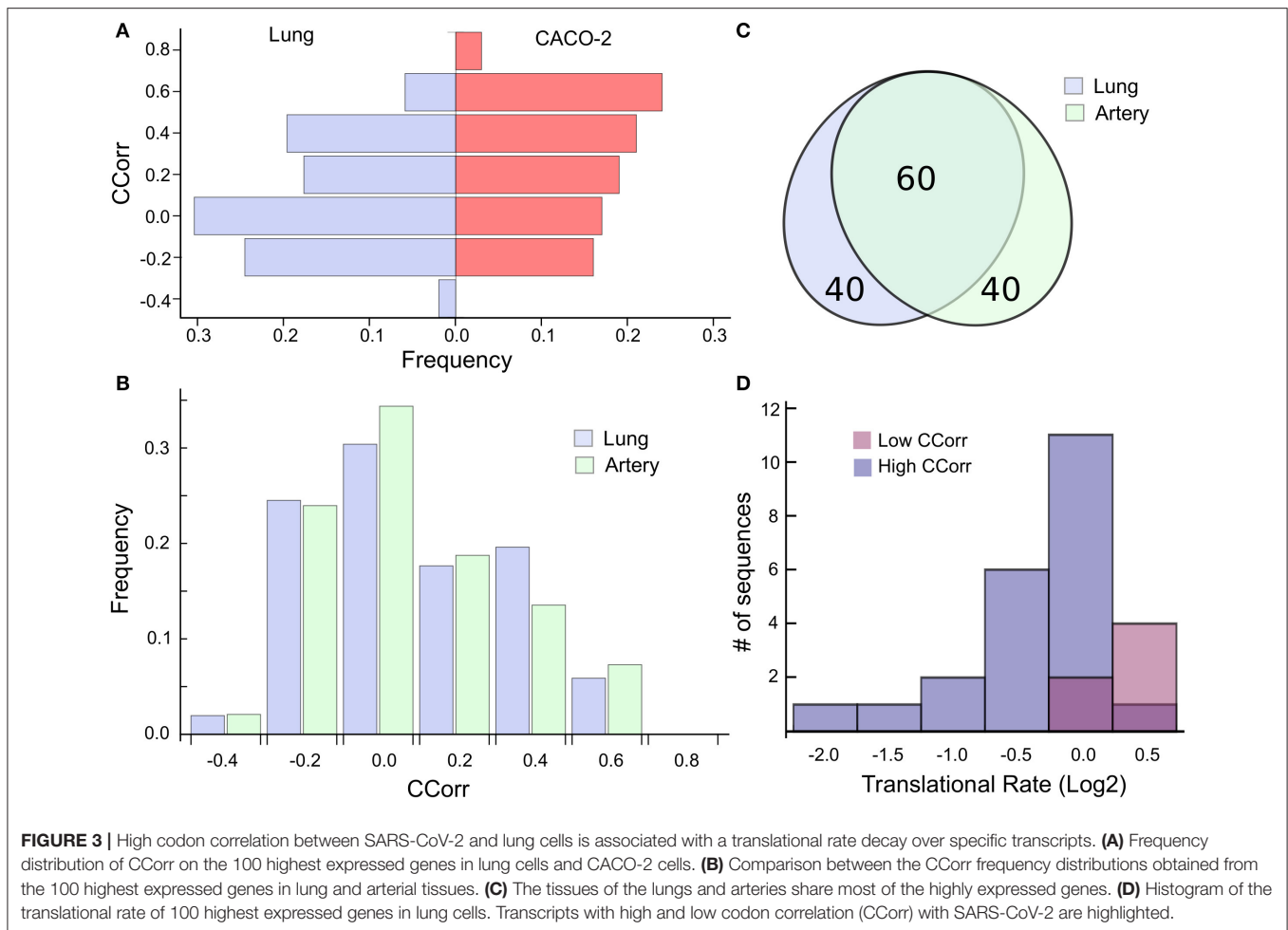
sequence in SARS-CoV-2-infected cells is lower than in mock-infected cells. Although the correlation ($\rho = -0.26$) with codon composition is not high, reflecting the fact that other factors could be regulating the translation rate, it is significant at a value of 0.05 (p-value $= 0.024 < 0.05$).

This analysis suggests then that the codon composition of highly expressed host-genes is a determinant of its own translation rate and that viral replication induces a particular case of epistasis, which could affect host cell proteostasis. Consequently, we find it relevant to extrapolate this phenomenon to cells affected by SARS-CoV-2. In this sense, the ACE2 receptor has been identified as the SARS-CoV-2 cell entry. Even though ACE2 transcript is present in almost all organs, the surface

expression of ACE2 protein is present in lung alveolar epithelial cells and arterial and venous endothelial cells (Hamming et al., 2004). For that reason, we elected to analyze the most expressed genes in lung and arterial tissues extracted from the GTEx database, two main targets of the SARS-CoV-2 infection. To this end, we select the 100 most expressed genes in these tissues and compute the CCorr for each sequence. **Figure 3A** shows that the codon composition profile of the most expressed genes from both lung and CACO-2 cells types are different, it is evident that the lung cells share fewer codons in common with the SARS-CoV-2 than the CACO-2 cells. On the other hand, the codon composition profile of the highly expressed genes in arterial tissue are quite similar to the genes associated with the lung (**Figure 3B**). In fact, they share 60% of these highly expressed genes (**Figure 3C**). Consequently, the tRNAs pool used for the virus could be scarcer both in lung and arterial cells than in the CACO-2 cells.

We search for those genes, highly expressed in the lung tissue, which could be affected by the depletion of the tRNAs consumed by virus replication. They are those whose codon composition is similar to SARS-CoV-2 (CCorr > 0.25); these are listed in **Table 1**. Before analyzing these 27 genes of interest, we compared the translation rate of these genes with the translation rate associated with highly expressed genes, but which differs in their codon usage bias (CCorr < −0.075). In this sense, we used the Mann-Whitney U test for median differences of independent samples to analyze the difference in translational rate between these two groups of genes. This test identified significant differences between these groups (p-value $= 0.0007$), as it can be seen in **Figure 3D**. We want to remark that, since there is still no available data on translational rate from lung cells, the last analysis was performed by using translational rate data available from CACO-2 cells, although we know that this data would be underestimating the difference between these two groups of genes.

FIGURE 3 | High codon correlation between SARS-CoV-2 and lung cells is associated with a translational rate decay over specific transcripts. **(A)** Frequency distribution of CCorr on the 100 highest expressed genes in lung cells and CACO-2 cells. **(B)** Comparison between the CCorr frequency distributions obtained from the 100 highest expressed genes in lung and arterial tissues. **(C)** The tissues of the lungs and arteries share most of the highly expressed genes. **(D)** Histogram of the translational rate of 100 highest expressed genes in lung cells. Transcripts with high and low codon correlation (CCorr) with SARS-CoV-2 are highlighted.

Furthermore, we performed a GO term enrichment analysis over the 27 genes listed in **Table 1**; their translation rates are decreased by means of the Enrichr online software (Chen et al., 2013). **Figure 4** illustrates the main enrichment pathways. This analysis reveals that codon usage could promote extensive changes in the translation machinery of the host in agreement with previous report in CACO-2 cells (Bojkova et al., 2020). It is known that when canonical translation is impaired, as part of the host defense program, specific 40S ribosomal subunits are needed to support uncapped viral mRNA translation (Kwan and Thompson, 2019).

Our list of genes from lung is enriched with several ribosomal proteins that constitute both 40S (RPS6, RPS3A, RPS7, RPSA, RPS25, RPS13, RPS12, and RPS27A) and 60S (RPL4, RPL5, RPL21, RPLP1, RPL34, RPL9, RPL24, and RPL17) ribosomal subunits. Many of them also appear in the set of genes derived from arterial tissue. All of them belong to the nonsense-mediated decay pathway that control the mRNAs with abnormal termination. The termination codon can be recognized if the 3′ untranslated region is short or if it does not have an exon junction complex downstream of the termination codon (Nicholson et al., 2010). These genes are also part of the process of SRP-dependent co-translational protein destined for the endoplasmic reticulum.

It is known that these proteins are involved in translation particularly the gene EEF1A1, an isoform of the alpha subunit of the elongation factor-1 expressed in lungs and arteries, which is responsible for the GTP-dependent binding of aminoacyl-tRNA to the ribosome. Beyond its function in translation, EEF1A1 takes part in the innate immune system by activating directly the transcription of IFN-gamma (Maruyama et al., 2007). IFN-gamma activates immune cells, such as macrophages and natural killer cells, and stimulates the major histocompatibility complex (MHC) II-dependent presentation of antigens to the immune system (Schroder et al., 2004). This process could be downregulated due to a decrease in the translation rate of EEF1A1. Furthermore, B2M, another gene in the list of decreased translational rate, encodes one of the proteins that conform MHC class I found on the cell surface of all nucleated cells (Bernier, 1980). Together with TXNIP, RPS3A, and RPS27A, it is involved in the antigen processing and presentation. Moreover, this gene, along with S100A11, HSP90AB1, and also EEF1A1, regulates the exocytosis of granules containing inflammatory mediators in neutrophils (Lee et al., 2005). All these genes, with the exception of TXNIP and S100A11, are also highly expressed in the arterial tissue. We highlight the presence in our analysis of the HSP90AB1, a known chaperone that facilitates the

TABLE 1 | Set of most expressed genes in lung (GTEx) and positive correlation with virus codon usage.

Gene name	Codon corr	T. rate 6 h	T. rate 10 h
RPS3A	0.582	0.686	0.980
RPL5	0.559	0.585	0.271
RPL9	0.546	0.723	0.738
SPARCL1	0.536	NA	NA
EEF1A1	0.531	0.948	0.627
RPS12	0.443	0.155	0.649
RPL17	0.438	NA	0.372
TPT1	0.415	1.048	0.981
RPL21	0.411	NA	NA
RPS6	0.398	1.206	0.590
RPL34	0.376	0.981	0.727
RPS13	0.372	1.386	0.209
RPL4	0.370	0.919	0.799
RPS27A	0.365	NA	NA
RPSA	0.352	0.902	0.643
SAT1	0.350	NA	NA
A2M	0.340	NA	NA
TXNIP	0.317	NA	NA
RPS25	0.315	0.728	0.753
FN1	0.308	NA	NA
RPL24	0.306	0.576	0.890
RPLP1	0.303	NA	NA
RPS7	0.297	0.965	1.036
HSP90AB1	0.296	0.989	0.738
B2M	0.277	0.926	0.825
S100A11	0.273	1.620	0.520
MGP	0.272	NA	NA

col 1: gene name, col 2: CCorr, cols 3 and 4: translational rate, infected CACO-2 cells, measured at 6 and 10 h PI, respectively. Translational rates are relative to mock cells.

maturation of a wide range of proteins and its attenuation has been related to idiopathic pulmonary fibrosis and cystic fibrosis (Haase and Fitze, 2015; Wang and Ni, 2016). Our analysis also shows a translation decrease in A2M, a plasmatic protein highly expressed in lung and artery that inhibits a broad spectrum of proteases, including trypsin, thrombin, and collagenase (Cater et al., 2019). A recent study confirms our prediction, A2M level in sera from COVID-19 patients is significantly lower than in healthy subjects (Shen et al., 2020). Furthermore, A2M has a key role in regulating inflammatory processes because is able to bind to proinflammatory ligands (Feige et al., 1996). This is particularly relevant in the context of the COVID-19-derived cytokine storm.

In addition, we have also identified two genes (SAT1 and MGP) belonging to the pathway of endothelial cell calcification regulated by NOTCH1 (White et al., 2015). This finding acquires special relevance in the context of the acute lung injury (ALI) observed in many infected patients (Huang et al., 2020). The first gene SAT1 catalyzes the acetylation of polyamines (spermidine and spermine) and carries it out of the cell. The polyamine excess is a prominent source of oxidative stress that can increase

inflammatory response (Hussain et al., 2017). Polyamines have also been connected with the immune system (Pérez-Cano et al., 2003). On the other hand, the MGP gene encodes the matrix gla protein, which is also highly expressed in all vasculatures. Recent studies suggest that MGP downregulates the tissue calcification by sequestering bone morphogenetic proteins (White et al., 2015). Mutations in this gene cause Keutel syndrome, which is characterized by peripheral pulmonary stenosis, abnormal cartilage calcification, and skin rashes (Munroe et al., 1999).

Another interesting protein predicted to be downregulated during the infection in lung and arterial tissues is SPARC-like 1 (SPARCL-1), also known as Hevin, commonly associated with regulation of cell migration and modulation of extracellular matrix proteins (Girard and Springer, 1996), and it has been shown to be involved in lymphocyte transendothelial migration through high endothelial venules (Girard and Springer, 1995). In this context, it is important to mention that clinical studies over patients that suffer several cases of COVID-19 document a dysregulation of immune response related particularly to a lower lymphocytes count (Huang et al., 2020). In addition, our analysis reveals a translational rate decay of fibronectin (FN1), a master organizer of extracellular matrices that mediates cellular interactions playing important roles in cell adhesion, hemostasis, and thrombosis (Pankov and Yamada, 2002; Wang and Ni, 2016). This prediction is in agreement with previous data showing that cells infected with SARS-CoV undergo downregulation in fibronectin expression (Surjit et al., 2004).

4. DISCUSSION

The viral infection of human cells triggers an ensemble of host processes based on the interferons that interfere with viral replication. These processes have co-evolved with the viral response to the host defense, and the virus counteracts these processes through a diversity of immuno-modulatory mechanisms. For example, NS1 protein plays a central role in the influenza infection by suppressing the host IFNs response. Recent results suggest that NS1 protein can hamper the host gene expression at the translational level by obstructing the mRNA entrance tunnel of ribosomes (Thoms et al., 2020). Furthermore, the N protein of porcine reproductive and respiratory syndrome virus impairs the IFN transcription by acting over the TXK, which together with the EEF1A1 and PARP1 form the trimolecular complex that binds to the IFN-γ gene promoter (Maruyama et al., 2007; Kenney and Meng, 2015). Hemagglutinin of IAVs has been shown to facilitate IFNAR ubiquitination and degradation, reducing the levels of IFNAR, and thus suppressing the expression of IFN-stimulated antiviral proteins (Xia et al., 2016). These examples illustrate the action of viral dedicated factors that downregulate the transcription of IFNs. Up to the present, no dedicated factor with analog function has been identified in SARS-CoV and SARS-CoV-2. However, a recent report found a significant lack of IFN type I and III at the transcriptional level in human alveolar adenocarcinoma cells (Blanco-Melo et al., 2020). On the other hand, a marked upregulation of inflammatory
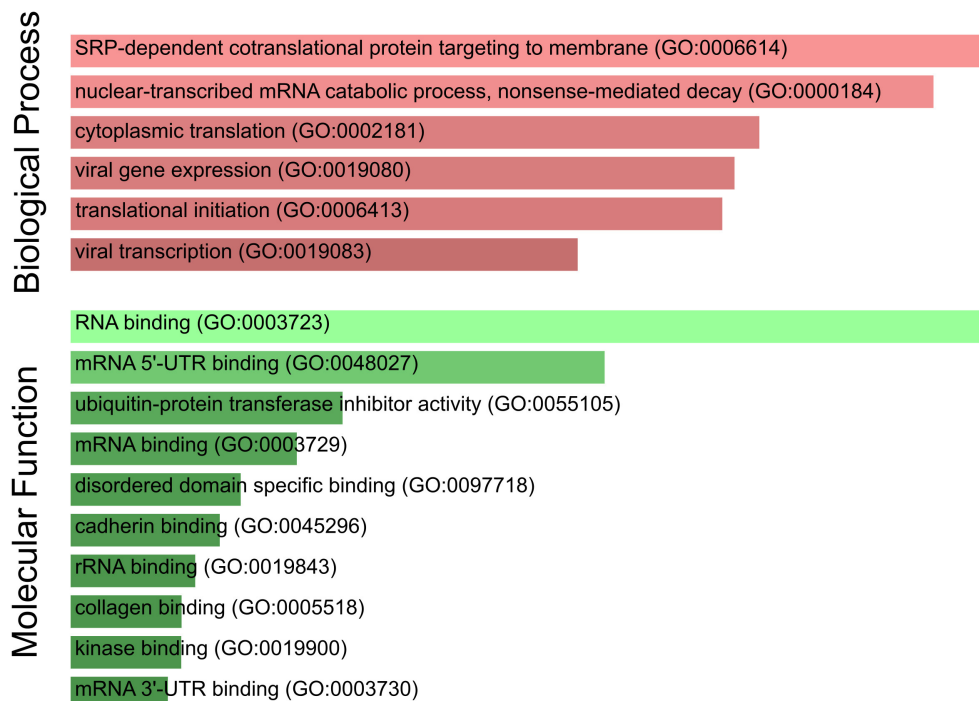
FIGURE 4 | SARS-CoV-2 codon usage may have an impact on host translational machinery. GO enrichment analysis performed over genes of interest. The most enriched biological processes and molecular functions are highlighted. Both size and color shade of bars represent the level of significance.

mediators at the protein level (CXCL10, CCL2, IFN-α, and γ) has been observed in patients with SARS-CoV without a significant amount of specific antibodies (Cameron et al., 2007). Several cases of absence of protective immunity due to previous infection seem to indicate a similar landscape for COVID-19. Until now, the manner in which some patients fail to develop adaptive immunity is yet to be elucidated. As mentioned in the Results section, the decreased translation of B2M could be related to this last observation since is a crucial factor for the stable presentation of antigens derived from virus or tumor proteins; these antigens are recognized by cytotoxic T cells that eventually eliminate the target cell stimulating apoptosis to prevent systemic dissemination of the disease (Hulpke and Tampé, 2013).

MGP is also expressed at high levels in heart, kidney, and lung which is particularly interesting in the context of several comorbidities and collateral effects observed in the COVID-19 patients (Nikolich-Zugich et al., 2020). For example, skin rashes were recently reported as a new symptom of COVID-19 and the authors postulate that recognizing rashes is important to identify new and earlier COVID-19, cases (Bataille et al., 2020). In this context, we highlight that permanent skin rashes were reported as a characteristic of the Keutel syndrome, and mutations in the MGP gene were reported as a crucial factor in this syndrome (Munroe et al., 1999; Khosroshahi et al., 2014). In this manner, our results are related to these observations.

Summing up, if the depletion of a selected set of tRNA, induced by virus replication, affects the expression level or the co-translation folding of these proteins, one could expect the emergence of several systemic disorders.

5. CONCLUSION

Codon usage bias is thought to have significant effects on translation rate, where rare codons are assumed to be translated more slowly than common codons (Piovesan et al., 2013). It is assumed that rare and common codons are defined by usage rates of highly expressed genes. However, whether the codon composition of viral ORFome can affect the translation rate of host genes has not been thoroughly explored yet. Here, we have shown that the synthesis of the proteins associated with highly expressed genes, and with similar codon usage to the one of the virus, appears to be downregulated. This finding is in agreement with recent observations in totivirus-infected yeast (Chen et al., 2020). Following this idea, we determined which genes in lung could be affected by the viral replication. A functional analysis of these genes reveals that they could be related to collateral effects observed in COVID-19 patients (Huang et al., 2020). Further studies are mandatory to corroborate or discard the putative relationship established here.

One of the main obstacles in the recent development of vaccines has been the finding of increased infectivity observed to occur after immunizations with whole virus vaccines or complete spike protein vaccines. This phenomenon has been observed both in vaccines against SARS coronavirus and in respiratory

syncytial virus. However, just as the virus regulates the translation of the host by its codon usage, the biotechnological manipulation of the frequency of codons could be used to design attenuated viruses. In this sense, other vaccine strategy has been recently assayed, focusing on altering the codon-pair usage without affecting protein sequence. This codon deoptimization strategy has reduced virus replication (Coleman et al., 2008; Le Nouen et al., 2014). We believe that our results shed light on how codon use could affect virus attenuation and would help decrease the damaging side effect, providing an exciting opportunity for live-attenuated vaccine development.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: http://corona.papers.biochem2.com; http://proteomecentral.proteomexchange.org/cgi/GetDataset? ID=PXD017710.

AUTHOR CONTRIBUTIONS

LD designed the experiments. AA prepared the data. LD and AA the analyzed the data and wrote the manuscript. All authors contributed to the article and approved the submitted version.

REFERENCES

Bataille, V., Visconti, A., Rossi, N., Murray, B., Bournot, A., Wolf, J., et al. (2020). Diagnostic value of skin manifestation of SARS-CoV-2 infection. *medRxiv*. doi: 10.1101/2020.07.10.20150656

Bernier, G. M. (1980). β2-microglobulin: structure, function and significance. *Vox Sanguinis* 38, 323–327. doi: 10.1159/000467018

Blanco-Melo, D., Nilsson-Payant, B. E., Liu, W. C., Uhl, S., Hoagland, D., Møller, R., et al. (2020). Imbalanced host response to SARS-CoV-2 drives development of COVID-19. *Cell* 318, E736–E741. doi: 10.1016/j.cell.2020.04.026

Bojkova, D., Klann, K., Koch, B., Widera, M., Krause, D., Ciesek, S., et al. (2020). Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature* 583, 469–472. doi: 10.1038/s41586-020-2332-7

Buhr, F., Jha, S., Thommen, M., Mittelstaet, J., Kutz, F., Schwalbe, H., et al. (2016). Synonymous codons direct cotranslational folding toward different protein conformations. *Mol. Cell* 61, 341–351. doi: 10.1016/j.molcel.2016.01.008

Cameron, M. J., Ran, L., Xu, L., Danesh, A., Bermejo-Martin, J. F., Cameron, C. M., et al. (2007). Interferon-mediated immunopathological events are associated with atypical innate and adaptive immune responses in patients with severe acute respiratory syndrome. *J. Virol.* 81, 8692–8706. doi: 10.1128/JVI.00527-07

Cater, J. H., Wilson, M. R., and Wyatt, A. R. (2019). Alpha-2-macroglobulin, a hypochlorite-regulated chaperone and immune system modulator. *Oxidat. Med. Cell. Longev.* 2019:5410657. doi: 10.1155/2019/5410657

Chen, E. Y., Tan, C. M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G. V., et al. (2013). Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* 14:128. doi: 10.1186/1471-2105-14-128

Chen, F., Wu, P., Deng, S., Zhang, H., Hou, Y., Hu, Z., et al. (2020). Dissimilation of synonymous codon usage bias in virus–host coevolution due to translational selection. *Nat. Ecol. Evol.* 4, 589–600. doi: 10.1038/s41559-020-1124-7

Coleman, J. R., Papamichail, D., Skiena, S., Futcher, B., Wimmer, E., and Mueller, S. (2008). Virus attenuation by genome-scale changes in codon pair bias. *Science* 320, 1784–1787. doi: 10.1126/science.1155761

Dana, A., and Tuller, T. (2014). The effect of tRNA levels on decoding times of mRNA codons. *Nucleic Acids Res.* 42, 9171–9181. doi: 10.1093/nar/gku646

Dhindsa, R. S., Copeland, B. R., Mustoe, A. M., and Goldstein, D. B. (2020). Natural selection shapes codon usage in the human genome. *Am. J. Hum. Genet.* 107, 83–95. doi: 10.1016/j.ajhg.2020.05.011

Diambra, L. A. (2017). Differential bicodon usage in lowly and highly abundant proteins. *PeerJ* 5:e3081. doi: 10.7717/peerj.3081

Feige, J.-J., Negoescu, A., Keramidas, M., Souchelnitskiy, S., and Chambaz, E. M. (1996). α2-macroglobulin: a binding protein for transforming growth factor-β and various cytokines. *Horm. Res. Paediatr.* 45, 227–232. doi: 10.1159/000184793

Frumkin, I., Lajoie, M. J., Gregg, C. J., Hornung, G., Church, G. M., and Pilpel, Y. (2018). Codon usage of highly expressed genes affects proteome-wide translation efficiency. *Proc. Natl. Acad. Sci. U.S.A.* 115, E4940–E4949. doi: 10.1073/pnas.1719375115

Gingold, H., and Pilpel, Y. (2011). Determinants of translation efficiency and accuracy. *Mol. Syst. Biol.* 7:481. doi: 10.1038/msb.2011.14

Girard, J. P., and Springer, T. A. (1995). Cloning from purified high endothelial venule cells of hevin, a close relative of the antiadhesive extracellular matrix protein SPARC. *Immunity* 2, 113–123. doi: 10.1016/1074-7613(95)90083-7

Girard, J. P., and Springer, T. A. (1996). Modulation of endothelial cell adhesion by hevin, an acidic protein associated with high endothelial venules. *J. Biol. Chem.* 271, 4511–4517. doi: 10.1074/jbc.271.8.4511

Gu, W., Zhou, T., Ma, J., Sun, X., and Lu, Z. (2004). Analysis of synonymous codon usage in SARS Coronavirus and other viruses in the Nidovirales. *Virus Res.* 101, 155–161. doi: 10.1016/j.virusres.2004.01.006

Haase, M., and Fitze, G. (2015). HSP90AB1: helping the good and the bad. *Gene* 575, 171–186. doi: 10.1016/j.gene.2015.08.063

Hamming, I., Timens, W., Bulthuis, M., Lely, A., Navis, G., and van Goor, H. (2004). Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J. Pathol.* 203, 631–637. doi: 10.1002/path.1570

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., et al. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395, 497–506. doi: 10.1016/S0140-6736(20)30183-5

Hulpke, S., and Tampé, R. (2013). The MHC I loading complex: a multitasking machinery in adaptive immunity. *Trends Biochem. Sci.* 38, 412–420. doi: 10.1016/j.tibs.2013.06.003

Hunt, R. C., Simhadri, V. L., Iandoli, M., Sauna, Z. E., and Kimchi-Sarfaty, C. (2014). Exposing synonymous mutations. *Trends Genet.* 30, 308–321. doi: 10.1016/j.tig.2014.04.006

Hussain, T., Tan, B., Ren, W., Rahu, N., Dad, R., Kalhoro, D. H., et al. (2017). Polyamines: therapeutic perspectives in oxidative stress and inflammatory diseases. *Amino Acids* 49, 1457–1468. doi: 10.1007/s00726-017-2447-9

Kandeel, M., Ibrahim, A., Fayez, M., and Al-Nazawi, M. (2020). From SARS and MERS CoVs to SARS–CoV–2: Moving toward more biased codon usage in viral structural and nonstructural genes. *J. Med. Virol.* 92, 660–666. doi: 10.1002/jmv.25754

Kenney, S. P., and Meng, X. J. (2015). An SH3 binding motif within the nucleocapsid protein of porcine reproductive and respiratory syndrome virus interacts with the host cellular signaling proteins STAMI, TXK, Fyn, Hck, and cortactin. *Virus Res.* 204, 31–39. doi: 10.1016/j.virusres.2015.04.004

Khosroshahi, H., Sahin, S., Akyuz, Y., and Ede, H. (2014). Long term follow-up of four patients with Keutel syndrome. *Am. J. Med. Genet. A* 164, 2849–2856. doi: 10.1002/ajmg.a.36699

Kim, A., Le Douce, J., Diab, F., Ferovova, M., Dubourg, C., Odent, S., et al. (2020). Synonymous variants in holoprosencephaly alter codon usage and impact the Sonic Hedgehog protein. *Brain* 143, 2027–2038. doi: 10.1093/brain/awaa152

Kirchner, S., Cai, Z., Rauscher, R., Kastelic, N., Anding, M., Czech, A., et al. (2017). Alteration of protein function by a silent polymorphism linked to tRNA abundance. *PLoS Biol.* 15:e2000779. doi: 10.1371/journal.pbio.2000779

Kwan, T., and Thompson, S. R. (2019). Noncanonical translation initiation in Eukaryotes. *Cold Spring Harbor Perspect. Biol.* 11:a032672. doi: 10.1101/cshperspect.a032672

Le Nouen, C., Brock, L. G., Luongo, C., McCarty, T., Yang, L., Mehedi, M., et al. (2014). Attenuation of human respiratory syncytial virus by genome-scale codon-pair deoptimization. *Proc. Natl. Acad. Sci. U.S.A.* 111, 13169–13174. doi: 10.1073/pnas.1411290111

Lee, K., Kang, H.-S., Jeon, J., Kim, E.-M., Yoon, S.-R., Song, H., et al. (2005). VDUP1 is required for the development of natural killer cells. *Immunity* 22, 195–208. doi: 10.1016/j.immuni.2004.12.012

Maruyama, T., Nara, K., Yoshikawa, H., and Suzuki, N. (2007). Txk, a member of the non-receptor tyrosine kinase of the Tec family, forms a complex with poly(ADP-ribose) polymerase 1 and elongation factor 1α and regulates interferon-γ gene transcription in Th1 cells. *Clin. Exp. Immunol.* 147, 164–175. doi: 10.1111/j.1365-2249.2006.03249.x

McCarthy, C., Carrea, A., and Diambra, L. (2017). Bicodon bias can determine the role of synonymous SNPs in human diseases. *BMC Genomics* 18:227. doi: 10.1186/s12864-017-3609-6

Munroe, P. B., Olgunturk, R. O., Fryns, J. P., Van Maldergem, L., Ziereisen, F., Yuksel, B., et al. (1999). Mutations in the gene encoding the human matrix Gla protein cause Keutel syndrome. *Nat. Genet.* 21, 142–144. doi: 10.1038/5102

Nicholson, P., Yepiskoposyan, H., Metze, S., Orozco, R. Z., Kleinschmidt, N., Mühlemann, O., et al. (2010). Nonsense-mediated mRNA decay in human cells: Mechanistic insights, functions beyond quality control and the double-life of NMD factors. *Cell. Mol. Life Sci.* 67, 677–700. doi: 10.1007/s00018-009-0177-1

Nikolich-Zugich, J., Knox, K. S., Rios, C. T., Natt, B., Bhattacharya, D., and Fain, M. J. (2020). SARS-CoV-2 and COVID-19 in older adults: what we may expect regarding pathogenesis, immune responses, and outcomes. *GeroScience* 42, 505–514. doi: 10.1007/s11357-020-00186-0

Pan, A., Dutta, C., and Das, J. (1998). Codon usage in highly expressed genes of *Haemophillus influenzae* and *Mycobacterium tuberculosis*: translational selection versus mutational bias. *Gene* 215, 405–413. doi: 10.1016/S0378-1119(98)00257-1

Pankov, R., and Yamada, K. M. (2002). Fibronectin at a glance. *J. Cell Sci.* 115, 3861–3863. doi: 10.1242/jcs.00059

Pérez-Cano, F. J., Franch, À., Castellote, C., and Castell, M. (2003). Immunomodulatory action of spermine and spermidine on NR8383 macrophage line in various culture conditions. *Cell. Immunol.* 226, 86–94. doi: 10.1016/j.cellimm.2003.09.009

Piovesan, A., Vitale, L., Pelleri, M. C., and Strippoli, P. (2013). Universal tight correlation of codon bias and pool of RNA codons (codonome): the genome is optimized to allow any distribution of gene expression values in the transcriptome from bacteria to humans. *Genomics* 101, 282–289. doi: 10.1016/j.ygeno.2013.02.009

Plotkin, J. B., and Kudla, G. (2011). Synonymous but not the same: the causes and consequences of codon bias. *Nat. Rev. Genet.* 12, 32–42. doi: 10.1038/nrg2899

Quax, T. E., Claassens, N. J., Söll, D., and van der Oost, J. (2015). Codon bias as a means to fine-tune gene expression. *Mol. Cell* 59, 149–161. doi: 10.1016/j.molcel.2015.05.035

Schroder, K., Hertzog, P. J., Ravasi, T., and Hume, D. A. (2004). Interferon-γ: an overview of signals, mechanisms and functions. *J. Leukoc. Biol.* 75, 163–189. doi: 10.1189/jlb.0603252

Shackelton, L. A., Parrish, C. R., and Holmes, E. C. (2006). Evolutionary basis of codon usage and nucleotide composition bias in vertebrate DNA viruses. *J. Mol. Evol.* 62, 551–563. doi: 10.1007/s00239-005-0221-1

Shah, P., and Gilchrist, M. A. (2011). Explaining complex codon usage patterns with selection for translational efficiency, mutation bias, and genetic drift. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10231–10236. doi: 10.1073/pnas.1016719108

Shen, B., Yi, X., Sun, Y., Bi, X., Du, J., Zhang, C., et al. (2020). Proteomic and metabolomic characterization of COVID-19 patient sera. *Cell* 182, 59–72. doi: 10.1016/j.cell.2020.05.032

Surjit, M., Liu, B., Jameel, S., Chow, V. T., and Lal, S. K. (2004). The SARS coronavirus nucleocapsid protein induces actin reorganization and apoptosis in COS-1 cells in the absence of growth factors. *Biochem. J.* 383, 13–18. doi: 10.1042/BJ20040984

Thoms, M., Buschauer, R., Ameismeier, M., Koepke, L., Denk, T., Hirschenberger, M., et al. (2020). Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. *Science* 17:eabc8665. doi: 10.1126/science.abc8665

Tsai, C.-J., Sauna, Z. E., Kimchi-Sarfaty, C., Ambudkar, S. V., Gottesman, M. M., and Nussinov, R. (2008). Synonymous mutations and ribosome stalling can lead to altered folding pathways and distinct minima. *J. Mol. Biol.* 383, 281–291. doi: 10.1016/j.jmb.2008.08.012

Tuller, T., Carmi, A., Vestsigian, K., Navon, S., Dorfan, Y., Zaborske, J., et al. (2010a). An evolutionarily conserved mechanism for controlling the efficiency of protein translation. *Cell* 141, 344–354. doi: 10.1016/j.cell.2010.03.031

Tuller, T., Waldman, Y. Y., Kupiec, M., and Ruppin, E. (2010b). Translation efficiency is determined by both codon bias and folding energy. *Proc. Natl. Acad. Sci. U.S.A.* 107, 3645–3650. doi: 10.1073/pnas.0909910107

von Herrmann, K. M., Salas, L. A., Martinez, E. M., Young, A. L., Howard, J. M., Feldman, M. S., et al. (2018). NLRP3 expression in mesencephalic neurons and characterization of a rare NLRP3 polymorphism associated with decreased risk of Parkinson's disease. *NPJ Parkinsons Dis.* 4:24. doi: 10.1038/s41531-018-0061-5

Wang, Y., and Ni, H. (2016). Fibronectin maintains the balance between hemostasis and thrombosis. *Cell. Mol. Life Sci.* 73, 3265–3277. doi: 10.1007/s00018-016-2225-y

White, M. P., Theodoris, C. V., Liu, L., Collins, W. J., Blue, K. W., Lee, J. H., et al. (2015). NOTCH1 regulates matrix gla protein and calcification gene networks in human valve endothelium. *J. Mol. Cell. Cardiol.* 84, 13–23. doi: 10.1016/j.yjmcc.2015.04.006

Xia, C., Vijayan, M., Pritzl, C. J., Fuchs, S. Y., McDermott, A. B., and Hahm, B. (2016). Hemagglutinin of influenza a virus antagonizes type I interferon (IFN) responses by inducing degradation of Type I IFN receptor 1. *J. Virol.* 90, 2403–2417. doi: 10.1128/JVI.02749-15

Zhou, T., Weems, M., and Wilke, C. O. (2009). Translationally optimal codons associate with structurally sensitive sites in proteins. *Mol. Biol. Evol.* 26, 1571–1580. doi: 10.1093/molbev/msp070